

High-throughput sequencing data reveals evolutionary conservation and differential transcription of satellites DNA among crickets species (Orthoptera; Gryllinae)

Palácios-Gimenez, O.M.^{1,2}; Bardella, V.B.¹; Lemos, B.²;
Cabral-de-Mello, D.C.¹

Abstract/Resumo

Satellite DNA (satDNA) is a class of non-coding repetitive DNA abundant in most eukaryotic genomes. Mostly satDNAs constitute clustered arrays of tandemly repeated sequences located in the gene-poor heterochromatin of centromeres and telomeres. Moreover in some cases, they are also dispersed in eu/heterochromatin of sex chromosomes and as single or short arrays nearby protein-coding genes within euchromatin. Here we take advantages of DNA- and RNA-seq data from cricket's species in order to perform a comparative analysis of content and expression of satDNAs. By graph-based clustering analysis of DNA-seq reads using RepeatExplorer software, dotplots analysis and FISH mapping we found that ~ 4% of the *G. assimilis* genome is represented by 13 A+T-rich satDNAs consisting of 11 well-defined families mainly located in the heterochromatic areas of chromosomes, and some of them able to form high-order repeats (HORs). In order to determine transcription profiles the raw RNA-seq reads from different tissues library of *Gryllus* species were mapped to each of the *G. assimilis* satDNAs using Bowtie2 and the method FPKM (fragments per kilo-base of transcript per million mapped reads) was used. The in silico transcriptional analysis of RNA-seq reads in *G. assimilis*, *G. bimaculatus*, *G. firmus* and *G. rubens* showed that some satDNAs are conserved in *Gryllus* species but differentially expressed in distinct tissues, sexes and besides tissue- and species-specific. In concordances with the transcriptional activity we found that *G. assimilis* satDNAs are also capable to adopt RNA secondary structures with well-defined helices ranging from 2 bp to 7 bp lengths. The folding possibility forming secondary structure helps to the satDNA dispersion along the genome by rolling-circle replication mechanism, in which circular monomer result from secondary structure RNA processing into linear monomers and subsequently circularization by a host-specific RNA ligase. The conservation of expression for different satDNAs in *Gryllus* species suggests a functional role for these sequences, as observed in other insects. Our data suggests functional roles of satDNAs for sexual differentiation at the chromatin levels, heterochromatin formation and centromeric function.

Keyword/Palavras-chave: Chromosomes; satDNA; FISH; DNA- RNA-seq

1 Departamento de Biologia, Instituto de Biociências/IB, Univ Estadual Paulista/UNESP, Rio Claro, São Paulo, Brazil, opalacios7@gmail.com
2 Molecular and Integrative Physiological Sciences Program, Department of Environmental Health, Harvard T. H. Chan School of Public Health, Boston, MA, USA